



# Variability in the Quality of Visual Working Memory

## Citation

Fougnie, Daryl, Jordan W. Suchow, and George A. Alvarez. 2012. Variability in the quality of visual working memory. *Nature communications* 3(11): 1229.

## Published Version

doi:10.1038/ncomms2237

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:10636292>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# Variability in the quality of visual working memory

Daryl Fougne<sup>1</sup>, Jordan W. Suchow<sup>1</sup>, & George A. Alvarez<sup>1</sup>

<sup>1</sup> Department of Psychology, Harvard University,  
Cambridge, Massachusetts 02138, USA

Correspondence and requests for materials should be addressed to D.F.  
(email: darylfougne@gmail.com).

## Abstract

Working memory is a mental storage system that keeps task-relevant information accessible for a brief span of time, and it is strikingly limited. Its limits differ substantially across people but are assumed to be fixed for a given person. Here, we show that there is substantial variability in the quality of working memory representations within an individual. This variability can be explained neither by fluctuations in attention or arousal over time, nor by uneven distribution of a limited mental commodity. Variability of this sort is inconsistent with the assumptions of the standard cognitive models of working memory capacity, including both slot- and resource-based models, and so we propose a new framework for understanding the limitations of working memory: a stochastic process of degradation that plays out independently across memories.

## Introduction

Working memory is critical to navigation, communication, problem solving, and other activities, but there are stark limits to how much we can remember<sup>1,9</sup>. A goal of research on working memory is to describe these limitations and determine their source.

Cognitive models typically explain the capacity of working memory by postulating a mental commodity that is divided among the items or events to be remembered<sup>1,3,5,9-12</sup>.

A core assumption of these models is that the quality (“fidelity”, “uncertainty”) of memories is determined solely by the amount of the commodity that is allocated to each item, being otherwise fixed<sup>1,5,10-13</sup>. Items receiving more of the commodity are better remembered.

A common approach to measuring the capacity of visual working memory is through the use of simple behavioral tasks. In one such task, participants are asked to remember the colors of a set of colorful dots, and then, after some delay, to report the color of a dot selected at random (Figure 1)<sup>1,9</sup>. Error is recorded as the difference between the correct color and the reported color. The distribution of these errors is

used to test models of memory. Standard models<sup>1,5,10-13</sup> assume that for each dot in the display, the participant is in one of two states, either remembering or forgetting. Importantly, these two states lead to different errors when the participant is asked to report the dot's color: errors for a forgotten item are uniformly distributed on the color wheel, whereas errors for a remembered item cluster around the correct color, with a spread that depends on how well it was remembered. The standard models assume that the form of the error distribution for remembered items is a circular normal (von Mises) distribution, with a spread that can be fully characterized by a single number: its *precision*, here defined as the standard deviation, as per [1].

(Figure 1 here)

It is possible to relax the assumption that the quality of memory is fixed, instead supposing that it varies within an individual. To determine whether such variability is present, we can inspect the shape of the error distribution. Critically, variability in the precision of memory within an individual leads to an error distribution that is more peaked than a circular normal distribution because it involves mixing error distributions that differ in precision but have the same mean (Figure 2). With sufficient data, it is possible to measure this peakedness to estimate the extent of high-order variability in the quality of memory within an individual.

(Figure 2 here)

Here, using variants of the behavioral task outlined above, we show that there is variability in the quality of working memory within an individual. This variability can be explained neither by fluctuations in attention or arousal over time, nor by uneven distribution of a limited mental commodity. We propose that a complete understanding of the capacity and architecture of working memory must consider both the allocation of a mental commodity and the stability of the information that is stored.

## **Results**

### **Revealing variability in the quality of working memory**

Participants were asked to perform a working memory task using displays with 1, 3, or 5 colorful dots (Figure 1). The standard fixed-precision model<sup>1</sup> and our variable-precision model was fit to each participant's distribution of errors (see Experimental Procedures). Figure 3 overlays the maximum a posteriori fixed- and variable-precision models on a histogram of errors made by one of the participants, for each of three set sizes: 1, 3, and 5. The variable-precision model provided a better fit to the data for each participant at each set size. When the two models were compared using the Akaike Information Criterion (AIC), a measure of goodness of fit, we found lower AIC values for the variable-precision model for each participant at each set size (Supplementary Table S1); the average magnitude of these AIC differences (7, 11.3, and 12) is a decisive victory for the variable precision model, leaving the fixed precision model with

“essentially no support” from the data<sup>14</sup>. Furthermore, the amount of variation in memory precision increased with set size for each participant (Figure 4). The average precision across subjects was 11.6°, 17.4°, and 23.7° for set sizes 1, 3, and 5, respectively. The average standard deviation of precision, a measure of variability in memory quality, was 3.0°, 5.4°, and 8.9°. These results were obtained with a variable-precision model in which the variability in precision was assumed to follow a truncated normal distribution. Evidence of variable quality was also found when substituting a gamma distribution over the inverse of the variance, similar to [15] (average inverse variance .009°, .005°, .003°; average standard deviation of inverse variance,  $2.3 \times 10^{-5}$ ,  $1.1 \times 10^{-5}$ ,  $9.0 \times 10^{-6}$ ).

(Figure 3 here)

(Figure 4 here)

### **Determining the source of the variability**

There are many possible sources of variability, including both true variability in memory precision as well as variability from other sources like perceptual noise, eye movements, and differences in the memorability of certain colors or locations. A series of control experiments were performed that demonstrate that these other sources made only a negligible contribution to the observed variability in memory quality.

When the memory demands of the task were eliminated by keeping the stimulus present while the participant responded, variability was eliminated and the data were better fit by a fixed-precision model (Supplementary Table S2). Even when the average

error on the perception task was increased to match that of the memory task (Supplementary Methods), variability was still three times smaller ( $2.2^\circ$  vs.  $8.9^\circ$ ,  $t(4) = 3.86$ ,  $p < .05$ , independent samples  $t$ -test).

A further experiment showed that there were negligible differences in the memorability of different colors (Supplementary Methods). We measured the variation in precision for six color bins, each spanning  $30^\circ$  of the color wheel ( $0-30^\circ$ ,  $60-90^\circ$ , ...,  $300-330^\circ$ ). Even though the range of colors within each bin was now only a small subset of the color wheel we still found substantial variability within a bin (Supplementary Figure 1). In fact, the variability within a bin was comparable to that in the full data set (Supplementary Figure 1c, one-sample  $t$ -tests,  $t(5)$ 's  $< .5$ ,  $p$ 's  $> .65$  for both participants). Furthermore, the variation in precision estimates across bins was minimal and accounts for at most a small (3.9%) portion of the total variance.

Similarly, we tested whether performance differed depending on the location of the tested item, a possible contributor to the observed variability in memory. For both the three- and five-item displays, we found little variance in the estimated precision across locations ( $4.15^\circ$  for 3-item displays,  $3.24^\circ$  for 5-item displays). (This is an upper bound on the contribution to variance because it includes measurement error.) Furthermore, the observed variance explains only 12.7% and 3.8% of the total variance for the 3-item and 5-item conditions, respectively.

Representational quality could differ for items that are fixated compared to items that are not. The encoding duration used in Experiment 1 was long enough to allow for a saccade, therefore raising the possibility that variability in precision is due entirely

to imbalances in encoding introduced by eye movements. We found that memories are variably precise even when eye movements are not possible. Two participants were asked to remember the color of three items presented for 100 ms. While this encoding duration is too short for an eye movement, we nonetheless found variability in the precision of working memory. Both participants' data were better fit by the variable- than fixed-precision model (average difference in  $AIC_c$  values of 4.0 in favor of the variable-precision model).

The previous manipulation eliminated the possibility of eye movements during stimulus presentation, but it does not address the possibility that gaze was not centered on fixation. We performed an additional follow-up experiment in which stimulus presentation was conditional on 1000 ms of continuous, uninterrupted fixation (Supplementary Methods). The data for two out of the three participants were better fit by the variable- than fixed-precision model, with an average difference in  $AIC_c$  of 6.7 in favor of the variable-precision model.

### **There is variability between items within a trial**

In the following experiments, we show that this variability can be explained neither by state-based fluctuations in attention or arousal across trials<sup>16</sup>, nor by the uneven allocation of a mental commodity within a trial. This pattern of results is not easily captured by models of working memory in which the quality of a memory is determined solely by the allocation of a finite commodity.



In a new experiment, on half of the trials participants were given the opportunity to report the color of the item they remembered best. On the other half of the trials, they reported the color of an item that was selected for them at random. These two conditions were intermixed in random order. Note that only variability within a trial can contribute to the participant's decision in this task. Therefore, if this variability is accessible to the participant, precision will be better for the best-remembered item than for one selected at random. This enables us to measure the extent to which the quality of memory varies within vs. across trials.

When probed on a random item, the mean precision was  $19.9 \pm 1.2^\circ$  (SEM) and its standard deviation was  $6.2 \pm 0.60^\circ$  (Figure 5a). However, when given the opportunity to choose which item to report, the mean precision was better ( $15.4 \pm .75^\circ$ , paired samples *t*-test,  $t(9) = 5.3$ ,  $p < .001$ ) and less variable ( $4.4 \pm .38^\circ$ ,  $t(9) = 3.4$ ,  $p = .01$ ) (Figure 5b).

This outcome provides an existence proof of within-trial variability and implies that it is cognitively present and useful. But how much of the variation is contained within a trial, and how much of it is across trials? To answer this question, we ran simulations under conditions of purely within- and purely across-trial variability (Figure 5c; see Experimental Procedures). We did not find a difference between the observed distribution and the simulated distribution for purely within-trial variation, either in mean ( $15.4^\circ$  observed vs.  $15.5 \pm 0.75^\circ$  simulated; paired samples *t*-test,  $t(9) = .05$ ,  $p > .85$ , ) or standard deviation ( $4.4^\circ$  observed vs.  $5.0^\circ \pm 0.41^\circ$  simulated;  $t(9) = 1.21$ ,  $p > .25$ ), (Figure 5d). Nor did we find a difference in the mean or standard deviation of

precision between the trials in which a random item was probed and trials in another control experiment, run with ten new participants, that did not include a “choose the best” condition ( $19.9^\circ$  vs.  $22.1 \pm 1.18^\circ$ , independent samples  $t$ -tests,  $t(18) = .91$ ,  $p = .38$ ;  $6.2^\circ$  vs.  $6.0 \pm .74^\circ$ ,  $t(18) = .74$ ,  $p = .48$ ). This suggests that giving the participant the opportunity to choose the best remembered item on some trials did not alter how items were encoded on the other trials.

(Figure 5 here)

In showing that nearly all of the variability in memory quality is accessible to the participant on a single trial, these results rule out a sizeable contribution of across-trial variability in the present study and suggest near-perfect metamemory for the relative quality of working memory representations. These findings also speak to an alternative account of the results of the first experiment, where we found that the distribution of response error was more peaked than is predicted by standard fixed-precision models<sup>1,5,10-13</sup>. Is it possible that the distribution’s peaked shape arose not because of variability in precision, but because of a fixed-precision error distribution that is naturally more peaked than a circular normal distribution? The results of the current experiment make this alternative account unlikely because nearly all of the peakedness in the distribution of report error is accounted for by variability across items within a trial, suggesting that the peakedness of the error distribution reflects variability across items and leaving little reason to posit an additional source.

### **Variability in memory quality is independent across items**

Many theories hold that the quality of a memory is determined by the allocation of a finite commodity, where items receiving more of the commodity are better remembered<sup>1,3,5,9,12</sup>. These theories predict that some of the within-trial variation in precision can be accounted for by uneven allocation: when more of the commodity has been allocated to one item, less is available for the others, thereby producing tradeoffs between items within a display. To determine whether the variability we observed was due to uneven commodity allocation, we performed an experiment in which participants were asked to report the color of every item in the display. We then tested for tradeoffs by comparing the errors made for one item when the responses to another item was good versus bad (i.e., above versus below the median absolute error for that response across trials) (Figure 6). We found no difference in precision between above-median and below-median splits for any combination of items, regardless of whether the estimates of precision were derived from the fixed-precision model (Figure 7a) (paired samples *t*-tests, all  $t(15)$ 's  $< 1.6$  and all  $p$ 's  $> .1$ ) or the variable-precision model (Figure 7b) (all  $t(15)$ 's  $< 1.2$  and all  $p$ 's  $> .25$ ).

(Figure 6 here)

(Figure 7 here)

This evidence of independence is based on a null result. However, using Monte Carlo simulations we found that our tradeoff-detection method is powerful enough to reliably detect tradeoffs if they were to exist (Supplementary Note). We considered two

sources of tradeoffs: (1) noiseless uneven allocation, where the participant knowingly attends or gives preferential processing to a chosen item, and (2) noisy even allocation, where the participant tries to distribute her resources evenly, but fails, unwittingly giving preference to some items over others. We found that any unevenness strong enough to produce the observed variability in precision would have been readily detected through our procedure for detecting tradeoffs (Supplementary Note). Finding independence of precision for items within a display<sup>17,18</sup> rules out that the observed variability in precision reflects unevenness in the allocation of discrete slots<sup>1</sup> or a graded resource<sup>5</sup>.

## **Discussion**

Visual working memory is not perfect. Its imperfections have been used to fashion cognitive models of visual working memory in which the quality of a memory is determined solely by the amount of a mental commodity that it receives. In contrast to the predictions of these models, the present results show that even under conditions of even allocation, the quality of working memory varies independently across trials and items. This finding is consistent with a stochastic process that degrades memory for each individual item and plays out independently across them. The outcome of this process leaves each memory in its own state, with its own precision, which when combined across items, produces the characteristic peaked shape of report-error distributions. Thus, visual working memory is stochastic both at the level of a memory's content and at the level of its quality.

A physiological mechanism that might account for variability in the quality of working memory representations is cortical noise that impedes the sustained activation of the neural populations that code for the remembered information<sup>5,12,19-23</sup>. The precision of memory representations may correspond to the amount of signal drift over time and the extent to which noisy signals are pooled within a neural population<sup>1,5,9,20,24-26</sup>. The recruitment of larger neural populations will improve the precision of working memory representations through decreased signal drift or increased redundancy in coding. Critically, the probability that neurons within a population will self-sustain after stimulus offset is known to be affected by internal noise<sup>20-22</sup>. Under these circumstances, the precision of memory representations would be determined by the outcome of stochastic physiological processes operating independently across representations.

Existing models explain the source of limitations in working memory solely by reference to a finite capacity for storing information. Here, we have shown that a complete account of visual working memory must also consider the stability of stored information. To put forward one model capable of producing variability of the sort described here, consider a finite mental commodity that must be divided amongst the items to be remembered. As in existing models, this information limit can be formalized as a set of  $N$  independent samples that are allocated to the items being remembered<sup>1,9-12</sup>, with slots models setting  $N$  to approximately 3, and with continuous resource models considering the limit as  $N$  tends to infinity. Suppose that the samples are unstable and that each has some probability of surviving until the time of the test.

This leads to variability in memory quality of a specific form: a binomial distribution of samples per item. This random process, known as the *pure death process* in studies of population genetics, is one of many possible random processes that might degrade memory.

Our model includes a guess state, a catch-all for trials on which the participant guesses randomly, either because of lapses in memory or because of other hiccups such as blinks and slips of the hand. A recent paper<sup>15</sup> has suggested that allowing for variability in the quality of working memory obviates the need to include a guess state. Though we agree that there is risk of over-estimating the rate of guessing if variability is ignored, we found that even taking variability into account, a guess state still accounts for a sizeable proportion of trials [Supplementary Notes]. Moreover, the goal of the present work is not to argue that participants never guess, but to show that variability in the quality of working memory reflects a random process of degradation that plays out independently across memories.

Models of working memory that are constrained by the known properties of the brain uniformly propose that the maintenance of working memory representations involves stochastic processes that operate at all levels of processing<sup>21,27,28</sup>. Yet most cognitive models explain working memory solely by reference to the division of a mental commodity: a resource divided among stored items<sup>1,3-5,9,10,12</sup>. In these models, the quality of a memory representation is deterministic and based solely on the amount of the commodity allocated to it. Demonstrating independent, stochastic variation in memory quality within an individual requires a new framework that includes a role for

stochastic processes, helping to bridge the gap between biologically-plausible neural models on the one hand and cognitive models on the other.

## Methods

### *Participants*

Participants were between the ages of 18 and 28, had normal or corrected-to-normal vision, and received either \$10/hour or course credit. Participants whose testing ran over multiple days were given a bonus of \$10 per session after the first. The studies were performed in accordance with Harvard University regulations and approved by the Committee on the Use of Human Subjects in Research under the Institutional Review Board for the Faculty of Arts & Sciences.

### *Stimuli*

The stimulus was 1, 3, or 5 colorful dots arranged in a ring around a central fixation mark. The radius of each dot was  $0.4^\circ$  in visual angle and the radius of the ring was  $3.8^\circ$ . Each dot was randomly assigned one of 180 equally-spaced equiluminant colors drawn from a circle (radius  $59^\circ$ ) in the CIE  $L^*a^*b^*$  color space, centered at  $L=54$ ,  $a=18$ ,  $b=-8$ . The grouping strength of the colors was controlled by transforming the color values to vectors in polar space, selecting displays with the constraint that the magnitude of the mean vector of each display was equal to that expected of a randomly-sampled display.

### *Presentation*

Stimuli were rendered by a computer running MATLAB with the Psychophysics Toolbox<sup>29,30</sup>. The display's resolution was  $1920 \times 1200$  at 60 Hz, with a pixel density of 38 pixels/cm. The viewing distance was 60 cm. The background was gray, with a luminance of  $45.4 \text{ cd/m}^2$ .

### *Procedure*

On each trial, the stimulus was presented for 600 ms and then removed. After a 900 ms delay, a filled white circle appeared at the location of a randomly selected item and hollow circles appeared at the location of the other items. Participants were asked to report the probed item's color by selecting it from a response screen with the full color wheel,  $6.5^\circ$  in radius, centered around fixation. A black indicator line was placed at the outer edge of the response wheel at the position closest to the cursor. Once the participants moved the mouse, the filled circle's color was continuously updated to match the currently-selected color. Participants registered their choice by clicking a mouse. Responses were not speeded. Feedback was provided after each response by displaying onscreen the error in degrees.

### *Experiment 1 Methods*

Three participants performed three sessions, each lasting 1.5–2 hours. Each session used 1, 3, or 5 items, and their order (and therefore set size) was counterbalanced across participants. There were 800 trials per session for set sizes 1 and 3 and 720 trials



per session for set size 5.

### *Experiment 2 Methods*

Ten participants performed two 700-trial sessions with a set size of 3. On half of the trials, participants reported the item they remembered best. On the other half, participants reported the color of a randomly selected item. Probe types were intermixed in random order. On “choose the best” trials, instructions appeared at the top of the screen telling participants to report a color and then to select the location of the best-remembered item by clicking its location. On these trials, filled white circles appeared at the locations of the probed and non-probed items.

### *Experiment 3 Methods*

Sixteen participants performed one 500-trial session with a set size of 3. Instead of reporting only one item, participants were asked to report all three. Items were probed in random order. Feedback was given. Non-probed items, even those already reported, were indicated by hollow circles.

## **Modeling the error distribution**

### *General framework*

All of the models considered here can be seen as special cases of an *infinite scale mixture*, a general framework that describes error distributions with a fixed mean and a

precision (scale) that is sampled from some higher-order distribution, known as the “mixing distribution”.

#### *Fixed-precision model*

In the fixed-precision model, the participant is in one of two states for each item on the display. With probability  $1-g$ , she remembers the item and has some fixed amount of information about it. Limited memory leads to errors in recall, which are assumed to be distributed according to a von Mises distribution (the circular analogue of a normal distribution) centered at zero (or, perhaps, with some bias  $\mu$ ) and with spread determined by a concentration parameter  $\kappa$  that reflects the memory’s precision. The mixing distribution is thus the Dirac delta function at  $\kappa$ , a distribution concentrated at that one point. With probability  $g$ , she remembers nothing about the item and guesses randomly, producing errors distributed according to a circular normal. Together, the probability density function of this model is given by

$$f(x; g, \mu, \kappa) = (g) \frac{1}{360} + (1 - g) \phi(x; \mu, \kappa),$$

where  $\phi$  is the von Mises density, defined by

$$\phi(x; \mu, \kappa) = \frac{e^{\kappa \cos(x - \mu)}}{360 I_0(\kappa)},$$

where  $I_0$  is the modified Bessel function of order zero.

#### *Variable-precision model*

The variable-precision model is the same as the fixed-precision model, except that precision is distributed according to some higher-order distribution, e.g., a truncated normal or gamma distribution. Unless guided by a theory for the source of variability in precision, the choice of mixing distribution is arbitrary, but certain combinations of error distribution and mixing distribution are convenient. For example, when error is normally distributed about the correct value, and when precision (i.e., the inverse of variance) is gamma distributed, the resulting experimental data takes the form of a generalized Student's  $t$ -distribution. (This result generalizes to the case of distributions wrapped on the circle, which is useful for representing colors and orientations.) For guess rate  $g$ , bias  $\mu$ , and standard deviation  $\sigma$ , the probability density function of this model is given by

$$f(x; \mu, \sigma, g, v) = (g) \frac{1}{360} + (1 - g) \psi(x; \mu, \sigma, v),$$

where  $\psi$  is the wrapped generalized Student's  $t$ -distribution<sup>31</sup> and is given by

$$\psi(x; \mu, \sigma, v) = \frac{c}{\sigma} \sum_{\ell=-\infty}^{\infty} \left( 1 + \frac{(x + 360\ell - \mu)^2}{\sigma^2 v} \right)^{-(v+1)/2},$$

with  $c = \Gamma\left(\frac{v+1}{2}\right) / \left(\Gamma\left(\frac{v}{2}\right) \sqrt{\pi v}\right)$ .

Incorporating higher order variability into new models of working memory is therefore a simple matter of replacing the usual choice of error distribution, a von Mises distribution, with a wrapped generalized Student's  $t$  distribution.

Some combinations of error and mixing distributions give rise to named distributions with known analytic expressions<sup>32,33</sup>, but others, like the truncated normal mixing distribution, must be simulated. The mean precision and its standard

deviation were free parameters, bounded between  $0^\circ$  and  $100^\circ$ . (The upper bound is arbitrary, but standard deviations above  $100^\circ$  are typically indistinguishable from guesses anyways.)

This should not be taken as a strong claim about the form of the distribution of memory precision, which will depend crucially on the process that degrades memories. However, no matter which distribution is chosen, the presence of variability produces a signature effect on the data — peakedness in the error distribution — seen most clearly in Figure 2, and which is readily detected using the present methods.

### *Model fitting*

Markov Chain Monte Carlo (MCMC) was used to find the maximum a posteriori parameter set of a model given the errors on the memory task. Each model was fit separately for each subject and experimental condition. The data were not binned. We used the Metropolis-Hastings algorithm, which takes a random walk over the parameter space, sampling locations in proportion to how well they describe the data and match prior beliefs. A non-informative Jeffreys prior was placed over each parameter. The proposal distribution used to recommend jumps was a multivariate Gaussian centered at the current parameter set, whose standard deviation was tuned during a burn-in period that ended when all of the chains, which started at different locations in the parameter space, converged. We collected 15,000 samples from these converged chains and report the sample with the maximum posterior probability.

### *Model comparison*

To compare the relative goodness of fit of the fixed- and variable-precision models, we computed each model's Akaike information criterion, a measure that includes a penalty term for the number of free parameters<sup>34</sup>. For a model with  $k$  parameters and maximum log-likelihood  $L$ , the Akaike information criterion is given by

$$\text{AIC} = -2L + 2k \quad (1)$$

Because this formulation of the AIC assumes infinite data, it is necessary to correct the value when estimating it from experimental data. The corrected AIC,  $\text{AIC}_c$ , is given by

$$\text{AIC}_c = \text{AIC} + \frac{2k(k+1)}{n-k-1}, \quad (2)$$

where  $n$  is the sample size—i.e., the number of trials performed by a participant on the working memory task.

### *Simulating across- and within-trial variation (Experiment 2)*

We performed Monte Carlo simulations to estimate the distribution of precision when 'choosing the best' under conditions of purely across-trial variation or purely within-trial variation. For each subject, 1,000 trials (of 3-item displays) were simulated according to that participant's maximum a posteriori parameter values for the random probe condition. When assuming across-trial variation, the precision value for each trial was selected at random from one of the three randomly drawn precision values. Error was simulated by sampling from a von Mises distribution with the drawn precision value. For each participant, the resulting error distributions were fit with the

variable-precision model, yielding parameter estimates that matched the random probe data (mean 20.2° and SD 6.4°; comparable to Figure 5, red line). When assuming within-trial variation, we selected the best precision value from the three that were randomly sampled. These values were used to simulate error responses that were then fit with the variable precision model (Figure 5, blue line, mean 15.4° and SD 5.0°).

## References

- 1 Zhang, W. & Luck, S. Discrete fixed-resolution representations in visual working memory. *Nature* **453**, 233-235, doi:10.1038/nature06860 (2008).
- 2 Vogel, E. K., Woodman, G. F. & Luck, S. J. Storage of features, conjunctions and objects in visual working memory. *Journal of Experimental Psychology: Human Perception & Performance* **27**, 92-114 (2001).
- 3 Alvarez, G. A. & Cavanagh, P. The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science* **15**, 106-111 (2004).
- 4 Luck, S. J. & Vogel, E. K. The capacity of visual working memory for features and conjunctions. *Nature* **390**, 279-281 (1997).
- 5 Bays, P. M. & Husain, M. Dynamic shifts of limited working memory resources in human vision. *Science* **321**, 851-854, doi:10.1126/science.1158023 (2008).
- 6 Rensink, R. A. The dynamic representation of scenes. *Visual Cognition* **7**, 17-42 (2000).
- 7 Baddeley, A. D. & Logie, R. in *Models of working memory: Mechanisms of active maintenance and executive control* (eds A. Miyake & P. Shah) 28-61 (Cambridge University Press, 1999).
- 8 Baddeley, A. D. *Working Memory*. (Oxford University Press, 1986).
- 9 Wilken, P. A detection theory account of change detection. *Journal of Vision* **4**, 1120-1135, doi:10.1167/4.12 (2004).
- 10 Anderson, D. E., Vogel, E. K. & Awh, E. Precision in Visual Working Memory Reaches a Stable Plateau When Individual Item Limits Are Exceeded. *The Journal of Neuroscience* **31**, 1128 -1138 (2011).
- 11 Zhang, W. & Luck, S. The Number and Quality of Representations in Working Memory. *Psychological Science* (2011).
- 12 Bays, P. M., Catalao, R. F. G. & Husain, M. The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision* **9**, 1-11 (2009).
- 13 Fougny, D., Asplund, C. L. & Marois, R. What are the units of storage in visual working memory? *Journal of Vision* **10**, 1-11 (2010).
- 14 Burnham, K. P. & Anderson, D. R. Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods and Research* **33**, 261-304 (2004).
- 15 Van den Berg, R. & Shin H., C. W. C., George, R. Ma, W.J. Variability in encoding precision accounts for visual short-term memory limitations. *PNAS* **109**, 8780-8785 (2012).
- 16 Kahneman, D. *Attention and effort*. (Prentice-Hall, 1973).
- 17 Fougny, D. & Alvarez, G. A. Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision* **11**, 1-12 (2011).
- 18 Huang, L. Visual working memory is better characterized as a distributed resource rather than discrete slots. *Journal of Vision* **10**, 1-8 (2010).

- 19 Deco, G., Edmund, T. R. & Romo, R. Stochastic dynamics as a principle of brain function. *Progress in Neurobiology* **88**, 1-16 (2009).
- 20 Camperi, M. & Wang, X. J. A model of visuospatial short-term memory in prefrontal cortex: recurrent network and cellular bistability. *Journal of Computational Neuroscience* **5**, 383-405 (1998).
- 21 Wang, X. J. Synaptic reverberation underlying mnemonic persistent activity. *Trends in Neurosciences* **24**, 455-463 (2001).
- 22 Constantinidis, C. & Wang, X. J. A neural circuit basis for spatial working memory. *The Neuroscientist* **10**, 553-565 (2004).
- 23 Fuster, J. & Alexander, G. Neuron activity related to short-term memory. *Science* **173**, 652-654 (1971).
- 24 Luce, D. R. Thurstone's discriminial processes fifty years later. *Psychometrika* **42**, 461-489 (1977).
- 25 Bonnel, A. M. & Miller, J. Attentional effects on concurrent psychophysical discriminations: Investigations of a sample-size model. *Attention, Perception, & Psychophysics* **55**, 162-179 (1994).
- 26 Thurstone, L. L. A law of comparative judgment. *Psychological Review* **34**, 273-286 (1927).
- 27 Durstewitz, D., Seamans, J. K. & Sejnowski, T. J. Neurocomputational models of working memory. *Nature Neuroscience* **3**, 1184-1191 (2000).
- 28 Durstewitz, D., Kelc, M. & Güntürkün, O. A Neurocomputational Theory of the Dopaminergic Modulation of Working Memory Functions. *The Journal of Neuroscience* **19**, 2807-2822 (1999).
- 29 Brainard, D. H. The Psychophysics Toolbox. *Spatial Vision* **10**, 433-436 (1997).
- 30 Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision* **10**, 437-442 (1997).
- 31 Pewsey, A., Lewis, T. & Jones, M. C. The wrapped t family of circular distributions. *Australian & New Zealand Journal of Statistics* **49**, 79-91 (2007).
- 32 Andrews, D. F. & Mallows, C. L. Scale mixtures of normal distributions. *Journal of Royal Statistical Society. Series B (Methodological)* **36**, 99-102 (1974).
- 33 Gneiting, T. Normal scale mixtures and dual propability densities. *Journal of Statistical Computation and Simulation*. **59**, 375-384 (1997).
- 34 Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**, 716-723 (1974).



**Acknowledgements:** We thank Sarah Cormeia and Christian Forhby for help with data collection. We are grateful to Geoff Woodman, Chris Asplund, Michael Cohen, Olivia Cheung, Justin Jungé, and Tim Brady for helpful discussions. This work was supported by NIH Grant 1F32EY020706 to D.F. and NIH Grant R03-086743 to G.A.A.

**Author contributions:** D.F., J.W.S, and G.A.A designed the experiments. D.F. performed the experiments. D.F. and J.W.S. analyzed the results. D.F., J.W.S, and G.A.A wrote the paper.

**Competing financial interests:** The authors declare no competing financial interests.

**Figure 1: Timeline of a working memory task.** First, a display of colorful dots is briefly presented. This is followed by an interval during which the participant is asked to hold the colors in mind. Finally, a response screen appears. On the response screen, the position of a randomly sampled item is cued and the participant is asked to select the color of the item that was presented at that location.

**Figure 2: Fixed and variable-precision models produce different distributions of error.** (a) Each curve shows the distribution of precision (the standard deviation of error on the memory task) for a different fixed-precision model. (b) A variable-precision model can be thought of as a higher-order distribution of fixed-precision models. One such higher-order distribution is shown here; it is a truncated normal with a mean of  $24^\circ$  and a standard deviation of  $7^\circ$ , bounded at  $0^\circ$  and  $100^\circ$ . (c) Variable precision models have different distributions of error. The green line is the distribution of errors for a model whose precision is fixed at  $24^\circ$ . (The green distribution in panels a and c are the same.) The red line is the expected distribution of errors if precision were drawn from the full distribution in panel b. Mixing together distributions with different precision produces a peaked shape that is a hallmark of the variable-precision model.

**Figure 3: Distinguishing between models with fixed- and variable-precision.** Each panel shows a histogram of errors on the working memory task and the best-fit fixed precision (green line) and variable-precision (red line) model. The working memory task used 1, 3, or 5 items (panels a, b, and c, respectively). These data are from a representative participant.

**Figure 4: Variability in the quality of working memory depends on the load.** The estimated distribution of precision according to the best-fitting variable-precision model for set sizes 1 (red), 3 (green), and 5 (blue), for each of three subjects (panels a–c). The bars along the horizontal axis show the precision values of the best-fit fixed-precision models.

**Figure 5: Choosing the item that was remembered best.** (a) When participants were asked to report an item selected at random, they made errors. (b) When allowed to report the item they remembered best, they performed better, seen as a narrower distribution of error. The ability to pick out the best-remembered item is evidence that the items differed in how well they were remembered. (c) Using the distribution in panel a, we simulated the distribution of errors that we would expect to find in panel b if participants had perfect metamemory, always picking out the best-remembered item. (d) The observed and simulated distributions of memory quality match, suggesting that participants have nearly perfect metamemory.

**Figure 6: Evidence that memory degradation is independent across items.** Distributions of report error for the first, second, and third responses sorted by whether the response to another item in the trial was below (black) or above (gray,

flipped) the median error. Memory quality does not differ between the two, which is evidence of memory degradation that is independent across items.

**Figure 7: Model fits for the data reported in Figure 6.** (a) We separately fit a fixed-precision model to the data from each participant, response (1st, 2nd, or 3rd), sorting, and split (above vs. below the median error). The best-fit standard deviation did not differ between the splits, which means that memory quality was comparable. (b) We did the same using the variable-precision model and found the same result: evidence of memory degradation that is independent across items.